ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue
Argonne, Illinois 60439

Propagation of data error and parametric sensitivity in computable general equilibrium model forecasts

Joshua Elliott, Meredith Franklin, Ian Foster, and Todd Munson

Mathematics and Computer Science Division

Preprint ANL/MCS-P1650-0709

July 2009, Revised March 2010

Contents

Propagation of data error and parametric sensitivity in computable general equilibrium model forecasts*

Joshua Elliott† Meredith Franklin‡ Ian Foster§ Todd Munson¶

March 26, 2010

Abstract

While the computable general equilibrium (CGE) model is a well established tool used in economic analyses, it is often viewed as a blackbox because of the complexity of the model structure and the many assumptions made regarding the underlying calibration data and model parameters. To characterize the behavior of the CGE model, we perform a large-scale Monte Carlo experiment to examine its sensitivity to two major forms of uncertainty: that caused by the expenditure data used to calibrate the model to a fixed base year and that resulting from the elasticity of substitution parameters at the core of the model. By examining a variety of output variables at different levels of economic and geographical aggregation, we assess how uncertainty impacts the conclusions that can be drawn from the model forecasts. We found greater sensitivity to uncertainty in the full set of elasticity of substitution parameters than to uncertainty in the base-year expenditure data as the forecast year increases. While many model forecasts were conducted to generate large output samples, we found that few forecasts are required to capture the mean model response of the variables we tested. However, characterizing the standard errors and empirical probability distribution functions was not possible without a large number of forecasts.

1 Introduction

The integrated assessment modeling (IAM) community couples economic and climate models to forecast the environmental and economic impacts of increasing greenhouse gas concentrations resulting from natural and anthropogenic emissions and to assess the effects of mitigation policies [5, 6, 8–10, 27, 29, 30]. Computable general equilibrium (CGE) models with nested constant elasticity of substitution (CES) production and utility functions have long been used by this community to model the economy. These models have two basic parameters: the share parameters and the elasticities of substitution among the commodities. The share parameters are derived from base-year expenditure data, while the elasticities of substitution are typically chosen by either expert opinion [31] or econometric estimates from time-series data [22]. All of these parameters are subject to uncertainty in their values, and the sensitivity of the models to this uncertainty is poorly

*This work was supported by grants from the MacArthur Foundation and the University of Chicago Energy Initiative and by the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy, under Contract DE-AC02-06CH11357.

†Computation Institute, U. Chicago and Argonne National Laboratory, email: jelliott@ci.uchicago.edu

‡U. Chicago and Argonne National Laboratory email: mfranklin@anl.gov

§Computation Institute, U. Chicago and Argonne National Laboratory, email: foster@anl.gov

¶Computation Institute, U. Chicago and Argonne National Laboratory, email: tmunson@mcs.anl.gov

understood. Moreover, failure to report information on the range of possible outcomes and their likelihood can result in misleading or erroneous conclusions.

Thus, computing the sensitivity in the forecasts generated by CGE models to parameter uncertainty is an essential task. Previous studies by the IAM community have been limited to a small subset of the parameters or to a simplified model having a single region. For example, Webster et al. [31] explored the sensitivity to the elasticities of substitution for capital, labor, and energy in their production functions and assumed these elasticities varied only by industry and remained constant across regions. Hertel et al. [21] looked at the sensitivity to only the Armington international trade elasticities. A study of the sensitivity to the share parameters derived from the base-year expenditure data has never been undertaken, likely because of the sample sizes necessary to study this large parameter space.

In this study, three large sets of model forecasts were conducted to examine the sensitivity of the CIM-EARTH CGE model [14] to uncertainty in the share parameters and the elasticities of substitution. Monte Carlo sampling of uncorrelated Gaussian distributions characterizing uncertainty in the parameters was performed to obtain the sample sets. Even though the parameter spaces we considered are large, we found that it is possible to obtain a robust understanding of these sensitivities with a reasonable number of samples. In Section 2, we review background on CGE models. In Section 3, we present the CIM-EARTH CGE model, detail the parameter distributions used for our analysis, and describe the statistical methods applied. In Section 4, we report the sensitivity results obtained from our generated forecasts. In Section 5, we summarize our findings.

2 Background

CGE models determine the market prices for commodities and their trade volumes such that supply equals demand [4,19,26]. Such models have many industries that each choose a feasible production schedule to maximize profit, many consumers that each choose a consumption schedule within their budget to maximize utility, and many markets where commodity and factor prices are set to clear the market. We use nested CES production functions to model the industry production constraints and nested CES utility functions for consumers.

Calibrated CES production functions [7] have the form

$$\mathbf{y} = \left(\sum_i \theta_i \mathbf{x}_i^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} ,$$

where $\mathbf{y}$ and $\mathbf{x}$ are the changes in output and inputs relative to a base year, respectively, and θ are the share parameters calibrated from base-year expenditure data with $\theta_i > 0$ and $\sum_i \theta_i = 1$. The elasticity of substitution parameter σ controls to what degree the inputs can be substituted for one another. When $\sigma = 0$, we obtain the Leontief production function

$$\mathbf{y} = \min_i \{\mathbf{x}_i\} ,$$

implying that the inputs are perfect complements; an increase in output requires an increase in *all* inputs. When $\sigma = 1$, we obtain the Cobb-Douglas function used extensively in economics:

$$\mathbf{y} = \prod_i \mathbf{x}_i^{\theta_i} .$$

The share parameter θ_i is the ratio of the base-year industry expenditure on that commodity to the total industry revenue. This choice of share parameters calibrates the model so that the output and inputs are consistent with the base-year expenditure data.

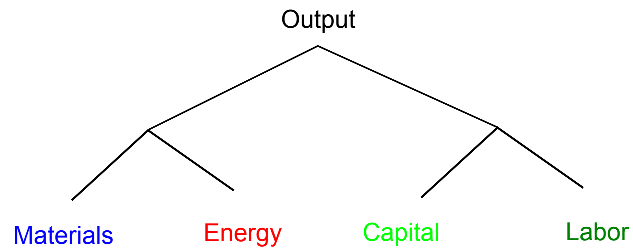

Figure 1: The tree structure for a simple production function.

These functions are then nested to constrain the decisions made by the industries and consumers. The structure is represented by a tree. The tree for a simple production function is shown in Figure 1. Each node represents a CES function with its own elasticity of substitution that aggregates the inputs from below into a commodity bundle. The root node aggregates the commodity bundles into the total industry output.

Because the optimization problems solved by the industries and consumers are convex in their own variables and satisfy a constraint qualification, we replace each with an equivalent complementarity problem obtained from the first-order optimality conditions by adding Lagrange multipliers on the constraints. These optimality conditions, in combination with the market-clearing conditions, form a square complementarity problem we can solve by applying the PATH algorithm [12, 15, 16], a Newton method designed to solve general square complementarity problems.

The resulting static model forecasts the equilibrium prices and quantities for a single year. Since we require forecasts for multiple years, we use a myopic dynamic CGE model to obtain multiyear forecasts in which the industries and consumers make optimal decisions based only on current information and do not consider the future. Specifically, we solve a sequence of static CGE models with exogenous dynamic trajectories for the factor endowments, efficiency units, and emission factors.

3 Methodology

The CIM-EARTH CGE model we use for this study is a myopic computable general equilibrium model with 16 regions; 16 domestic industries, 16 importers, and 1 capital goods industry per region; 1 consumer and 4 production factors per region; 3 homogenous transportation service industries; and 60 time periods (2004–2063). Capital in this model is perfectly mobile across industries, but not mobile across regions. We use exogenous dynamic trajectories for labor productivity and supply, energy efficiency, and resource availability based on the dynamic equations of the EPPA model [2]. Carbon emission amounts from fossil fuel usage are derived from the energy volume information in GTAP-E [11]. The CIM-EARTH CGE model is coded in the AMPL modeling language [17] and solved by applying the PATH algorithm [12, 15, 16]. For complete details, see the CIM-EARTH CGE model documentation [14].

The regions, industries, and factors in the CIM-EARTH CGE model are found in Table 1. Each region is labeled with our estimate of the level of uncertainty in the economic data: (L) low, (M) medium, and (H) high. For each industry we indicate the structure of the production functions: (A) agriculture, (E) extraction of fossil fuels, (M) manufacturing, (N) electricity generation, (P) petroleum refining, and (S) service industries.

The structure of the production functions for each industry type is loosely based on the EPPA

Table 1: Regions, industries, and factors for the CIM-EARTH CGE model. Regions are labeled with our estimate of the level of uncertainty in their economic data: (L) low, (M) medium, and (H) high. The industries are labeled by their production function structure: (A) agriculture, (E) extraction of fossil fuels, (M) manufacturing, (N) electricity generation, (P) petroleum refining, and (S) service industries.

Regions	**Industries**	**Factors**
Canada (L)	Agriculture and Forestry (A)	Capital
Mexico (M)	Coal Extraction (E)	Labor
United States (L)	Gas Extraction (E)	Land
Brazil (H)	Oil Extraction (E)	Natural Resources
Rest of Latin America (H)	Cement (M)	
Western Europe (L)	Chemicals (M)	
Rest of Europe (M)	Nonferrous Metals (M)	
Middle East and North Africa (M)	Steel and Iron (M)	
Rest of Africa (H)	Other Manufacturing (M)	
China, Mongolia, and Koreas (H)	Electricity (N)	
India (H)	Petroleum Refining (P)	
Japan (L)	Air Transport (S)	
Russia, Georgia, and Asiastan (M)	Land Transport (S)	
Rest of South Asia (H)	Sea Transport (S)	
Rest of Southeast Asia (M)	Government Services (S)	
Oceania (L)	Other Services (S)	

model [2] and summarized by the tree structure in Figure 2. The structure of the production functions for the importers of each commodity in each region is also provided. The capital goods industries aggregate materials using a single Leontief production function and do not demand fossil fuels, refined petroleum, electricity, or production factors; these capital goods are demanded only by consumers. The homogeneous transportation service industries simply aggregate air, land, and sea transportation services from each region into a single commodity using a Leontief production function; these homogeneous transportation services are used only for international trade.

The forecasts generated by this model are highly dependent on the choices of values for the share parameters and the elasticities of substitution, and thus on the data from which they are estimated. Since the share parameters are computed from the base-year expenditure data, we refer to uncertainty in this data as *expenditure uncertainty*. We refer to uncertainty in the elasticity of substitution parameters as *elasticity uncertainty*. To make our studies tractable, we limit the number of uncertain parameters by applying simplifying assumptions. We now document our methodology for studying the sensitivity of forecast variables to these uncertainties.

3.1 Expenditure Uncertainty

Since the share parameters are computed from base-year expenditure data, we can equivalently determine the sensitivity to uncertainty in the expenditure data. The CIM-EARTH CGE model uses the GTAP version 7 database for the 2004 base-year expenditures [20].[1] The primary source

[1]The GTAP database is constructed by removing inconsistencies in the raw expenditure data supplied for each region using a balancing routine. Ideally, we would treat the raw expenditure data as uncertain and apply this balancing routine to samples drawn from the uncertain raw data. However, using this process was not possible because of restricted access to both the raw expenditure data and the balancing routine for GTAP and to the large

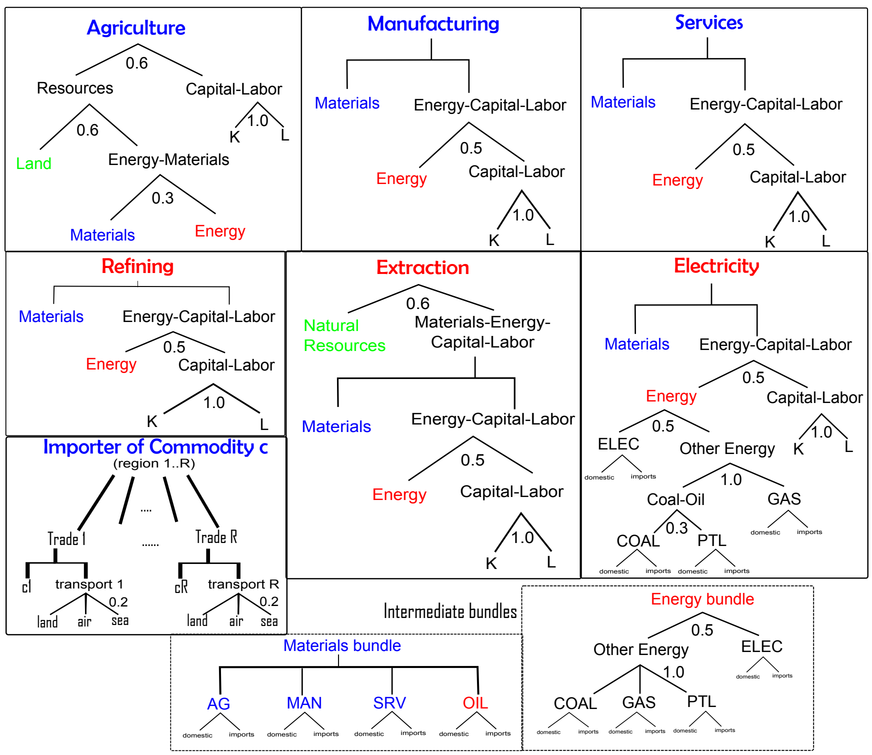

Figure 2: CIM-EARTH CGE nested CES production functions. Each node represents a CES function. Nodes with vertical line inputs represent Leontief functions; the other nodes are labeled with their elasticities of substitution.

of error in this database is reporting error. Another source of error, however, is the lack of updated data. For example, of the 87 regions and countries in the GTAP version 6 database with a 2001 base year, less than half reported new 2004 expenditure data for version 7. When new data is not reported, the 2001 expenditure data is scaled to obtain 2004 expenditure data; this 2001 expenditure data may in fact be a scaled version of 1997 expenditure data from GTAP version 5.

The full GTAP version 7 database has 113 regions (R), 57 industries (I), and 5 factors (F). The maximum number of expenditure parameters in a particular $R \times (I + F)$ aggregation is $RI(2I + F)$ for industries, $RI(4R)$ for importers (including land, sea, and air transportation), and $R(2I + 1)$ expenditures for consumers (including demand for savings). Therefore, the full database has a maximum of approximately 3.7 million expenditure values. Many of these expenditure values, however, are zero. With 16 regions, 16 industries, and 4 factors, the CIM-EARTH CGE model has a maximum of around 26,000 expenditure values. Approximately 40% of these values are ignored because the expenditure is less than $1M. The remaining parameter space, however, is still large. For this study, we chose a parameter space consisting of the 100 largest expenditure values per region, with a maximum of 5 expenditures chosen from each industry, leading to 1,600 uncertain expenditure parameters. While we restricted the analysis to this smaller set of expenditures, they still account for more than 75% of global expenditures.

We use uncorrelated Gaussian distributions to characterize the uncertainty in the expenditure values; taxes and subsidies on those expenditures were chosen to maintain constant rates.[2] The standard deviation of the distribution about the mean expenditure value is based on our estimate of the level of uncertainty in the economic data for each region. The regions indicated by (L) in Table 1 are believed to have well-established structures in place for consistent and accurate data gathering. Therefore, we assume the reported expenditure values from these regions have low levels of uncertainty and set the standard deviation to 3% of the mean value. In contrast, we assume that poorly developed regions and regions notorious for having data inconsistencies have high levels of uncertainty in their reported expenditure values. For these regions indicated by (H) in Table 1, we set the standard deviation of each distribution to 7% of the mean value. For all other regions indicated by (M) in Table 1, we assume medium levels of uncertainty and set the standard deviation of each distribution to 5% of the mean value. Unsurprisingly, the regions with low levels of uncertainty in their economic data have updated expenditure data for 2004 in the GTAP version 7 database. Most regions categorized as having a medium or high level of uncertainty in their economic data, including China, South and Southeast Asia, and most of Africa and Latin America, did not provide updated expenditure data for 2004. Rather, the 2004 expenditure data in the GTAP version 7 database was obtained by scaling the 2001 expenditure data from GTAP version 6.

3.2 Elasticity Uncertainty

The elasticities of substitution in our model are based on parameters from the EPPA [28, 31] and GTAP [25] CGE model documentation and on recent estimates derived from historical U.S. Bureau of Economic Analysis data [3]. The elasticities of substitution for most commodity bundles are

number of expenditure values in the fully disaggregated database.

[2]While some expenditure values may be correlated, we did not account for this possibility in the uncertainty distributions, even though such correlations would have an effect on the sensitivity of the forecast variables. Establishing the existence and extent of the correlation would require a more detailed examination of the underlying covariance structures. The actual tax and subsidy amounts and carbon emission factors may also be uncertain because of reporting errors or lack of updated data. We did not consider the sensitivity of forecast variables to uncertainty in these values.

Table 2: Mean elasticity of substitution parameters between domestic and imported commodities and the Armington international trade elasticities by industry for the CIM-EARTH CGE model.

	Elasticity of Substitution	
Industry	Domestic/Import	Armington
Agriculture and Forestry	2.7	5.6
Coal Extraction	3.0	6.1
Gas Extraction	17.2	34.4
Oil Extraction	5.2	10.4
Cement	2.9	5.8
Chemicals	3.3	6.6
Nonferrous Metals	4.2	8.4
Steel and Iron	3.0	5.9
Other Manufacturing	3.4	7.2
Electricity	2.8	5.6
Petroleum Refining	2.1	4.2
Air Transport	1.9	3.8
Land Transport	1.9	3.8
Sea Transport	1.9	3.8
Government Services	1.9	3.8
Other Services	1.9	3.8

indicated in Figure 2; the elasticities of substitution between domestic and imported commodities and the Armington international trade elasticities are found in Table 2.

The elasticity of substitution parameters play a key role in the forecasts generated by CGE models, yet there are vast inconsistencies in the estimates used by different models. The plot in Figure 3, for example, compares the estimates for the elasticity of substitution between capital and labor in the coal extraction industry used by the EPPA [28] and GTAP [25] CGE models and the estimate from Balistreri et al. [3]. Moreover, when we look at the elasticity of substitution between capital and labor across many industries, we can discern no consistent pattern in the disagreement. The GTAP parameter, for example, is at times the lowest and at other times the highest. This analysis is summarized in Appendix A.

For this study, we follow previous studies [21,28,31] and do not allow the elasticities of substitution to vary by region because of both a lack of data to support this differentiation and a desire to reduce the number of uncertain parameters. Therefore, the production functions for each industry have the same structure and elasticity of substitution parameters independent of the region. Moreover, we assume that the Leontief nests in the production functions are certain. Our assessment of uncertainty therefore consists of evaluating model forecast sensitivity to 70 elasticity of substitution parameters obtained by aggregating similar industries. The complete list of parameters is found below.

- 16 elasticity of substitution parameters between capital and labor
- 16 elasticity of substitution parameters between domestic and imported commodities
- 16 Armington international trade elasticities
- 5 additional elasticity of substitution parameters common across the (A) industries

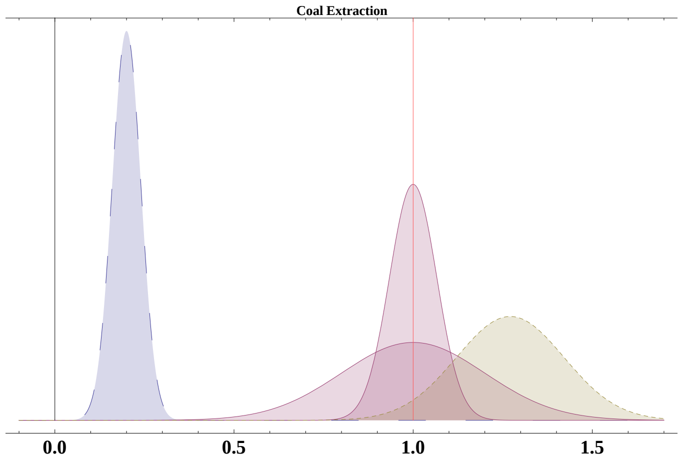

Figure 3: Comparison of distributions for the elasticity of substitution between capital and labor for the coal industry. Blue dashed: GTAP estimate with standard deviation set to 20% of the mean; yellow dotted: estimate from Balistreri et al. [3]; red solid: two distributions with Cobb-Douglas mean and standard deviation estimate from EPPA [28] and standard deviation set to 0.20.

- 4 additional elasticity of substitution parameters common across the (E) industries
- 3 additional elasticity of substitution parameters common across the (M) industries
- 4 additional elasticity of substitution parameters common across the (N) industries
- 3 additional elasticity of substitution parameters common across the (P) industries
- 3 additional elasticity of substitution parameters common across the (S) industries

We use uncorrelated Gaussian distributions to model the uncertainty in these elasticity of substitution values.[3] Instead of trying to combine disparate and often contradictory estimates of means and standard deviations, we centered our parameter distributions at relatively standard mean values and set the standard deviation to 20% of the mean.

A more detailed examination of the major sources of sensitivity was conducted by focusing specifically on the sensitivity to uncertainty in the 16 Armington international trade elasticities. Parameters for this study were again assumed to be the same in every region.

3.3 Statistical Methods

Approaches to quantifying model uncertainty include statistical sensitivity analysis [1, 18] and stochastic parameter control [23, 24]. The statistical approach evaluates the CGE model using random samples drawn from distributions defined around the uncertain parameters. Monte Carlo methods, for example, can be conveniently applied to sample the distributions but require significant computational power for large parameter spaces. Stochastic parameter control accounts for the temporal evolution of uncertain parameters and their relationships by incorporating estimated variance-covariance matrices.

[3] While some elasticity of substitution parameters may be correlated among similar industries, for example, we did not account for this possibility in the uncertainty distributions, even though such correlations would have an effect on the sensitivity of the forecast variables. Establishing the existence and extent of the correlation would require a more detailed examination of the underlying covariance structures.

We took the statistical approach using Monte Carlo methods to sample from the distributions of expenditures and elasticities described above. In particular, we constructed three large sample sets

- 10,000 samples drawn from our 1,600 uncertain base-year expenditure parameters
- 5,000 samples drawn from the 70 uncertain elasticity of substitution parameters
- 1,000 samples drawn from only the 16 uncertain Armington international trade elasticities

We used the Swift scripting system [32, 33] to evaluate the CIM-EARTH CGE model for each sample in parallel. We used roughly 30,000 CPU-hours to complete these forecasts; each forecast took 0.4-1.6 hours to complete. At the peak, approximately 2,000 simultaneous processors were used. For some samples, the PATH algorithm failed to compute a solution to within its default convergence tolerance. Those results were discarded, leaving 9,906 successful samples for the expenditure dataset, 4,978 successful samples for the elasticity dataset, and 999 successful samples for the Armington dataset.

We used bootstrap resampling [13] to explore the extent to which a smaller number of model forecasts would have sufficiently characterized the uncertainty in the forecast variables. In the context of this study, bootstrapping refers to iteratively choosing random subsamples of various sizes from the full set of model forecasts with replacement and then calculating statistics for the subsample. We wanted to determine the subsample size that provides a statistical approximation of the full set with 95% confidence. To this end, we performed several two-sample tests between the bootstrap statistics of the subsample and full set. Specifically, equality in the bootstrap means was assessed with a t-test; equality in the standard errors, the square root of the variance of the resampled statistic, was assessed with an F-test; and equality in the empirical probability distribution functions was assessed with the Kolmogorov-Smirnov KS-test. The p-value of these tests indicates whether to accept the null hypothesis that the two bootstrap samples are statistically equal. Thus, the p-value allows us to ascertain what subsample size replicates the full set with 95% confidence.

4 Results

We now present analysis of the results obtained from these forecasts. Our objective is to understand their sensitivity to expenditure and elasticity uncertainties. The variables we report are the gross domestic product, aggregate CO_2 emissions, revenue for the steel and iron industry, and aggregate industrial and consumer electricity demand for specific regions in the model and for a global aggregate. To summarize the forecast variables generated from the model forecasts, we report the coefficient of variation expressed as a percent ($\%c_v$); it provides a measure of the relative sensitivity to uncertainty. Specifically, it is the ratio of the sample standard deviation to the sample mean; larger values indicate greater model sensitivity.

When comparing forecasts of a particular variable taken at initial and terminal years, we report the sample correlation coefficient r, which measures how much of the uncertainty in two sample sets can be explained by a linear regression model. The percentage of the uncertainty that can be explained by the linear regression is $100r^2$. Thus, for high correlation coefficients, $r > 0.7$, the uncertainty in the variable at the base year explains approximately half of the uncertainty in the terminal year. We also report the slope β of the sample linear least-squares regression model. To first order, the slope describes the overall rate at which the distribution of the forecast variable changes from the base to terminal years.

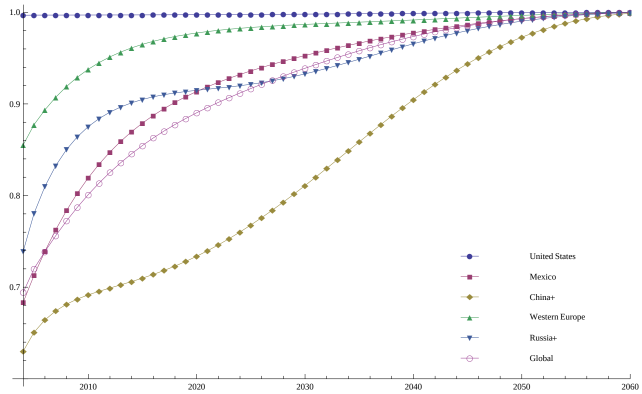

Figure 4: Correlation coefficient r between gross domestic product in the years 2004 to 2060 and the terminal forecast year, 2063.

4.1 Sensitivity to Expenditure Uncertainty

Table 3 summarizes the 2004 and 2063 coefficient of variation and linear correlation over the 9,906 viable forecasts for several variables. We find that global aggregations generally have smaller coefficients of variation than their regional equivalents for both the base and terminal years, suggesting that large-scale aggregates are less sensitive to uncertainty in the expenditure data than are regional aggregates. This observation is likely an effect of attenuation toward the sample average; aggregating to the global scale results in cancellation of regional forecast variability. When exploring the sample correlation coefficient between the years 2004 and 2063 of the forecast variable, we find that the global aggregates generally have much weaker correlations than do their regional counterparts. This result is opposite the behavior of the coefficient of variation. The growth in the coefficient of variation from the year 2004 to the year 2063 is significantly smaller for the (L) regions. Moreover, the sample correlation coefficient in the (L) regions is very high for all variables. Over 70% of the uncertainty in their 2063 values is explained by the uncertainty in their 2004 values. In the (H) and (M) regions, r is generally smaller, suggesting that the uncertainty in their 2063 values is not well explained by the uncertainty in their 2004 values. The evolution of the correlation coefficient over time depicted in Figure 4 shows that regions with high initial uncertainty, such as China, have a gradual, near-linear increase as the gap between the initial and forecast years narrows. Comparatively, Western Europe and the United States quickly plateau to a relatively high correlation coefficient, indicating that we can explain much of the variability in the terminal forecast year with earlier year forecasts.

We infer from these results that when the coefficient of variation is small, few simulations are necessary to characterize the model response with confidence. When correlation is large, we can predict the behavior of the trend and subsequently its uncertainty for that particular forecast variable. Thus, in this situation, fewer forecasts would be necessary to extract the same amount of information.

To explore this observation further, we took 1,000 bootstrap resamples for sample sizes ranging from 5 to 9,900 of one of the most sensitive variables in our model, 2063 Chinese consumer demand for electricity. The bootstrap mean and its standard error are plotted against the resample size in Figure 5. The results of the two sample t-tests indicate that the bootstrap mean is statistically equal to the full sample mean for all sample sizes. That is, the p-value is greater than 0.05, and we fail to reject the null hypothesis. Thus, we can recover the mean model response with as few as 5 model

Table 3: Coefficient of variation, sample correlation, and linear regression coefficients for select variables and regions over the 9,906 viable forecasts performed to assess sensitivity to expenditure uncertainty.

Variable	Region	$\%c_v$ 2004	$\%c_v$ 2063	r	β
Gross Domestic Product	Mexico	1.43	2.46	0.68	4.99
	China+	1.95	4.15	0.63	13.68
	Western Europe	1.16	1.80	0.86	4.95
	Russia+	1.51	2.78	0.74	10.70
	Global	0.69	0.85	0.69	4.21
Aggregate CO_2 Emissions	United States	1.41	1.96	0.88	3.33
	Mexico	3.08	4.13	0.88	3.39
	China+	2.00	3.65	0.49	5.31
	Western Europe	1.49	2.53	0.87	3.10
	Russia+	2.30	4.13	0.76	6.49
	Global	0.70	1.68	0.54	4.72
Steel and Iron Industry Revenue	United States	1.76	2.54	0.84	4.87
	Mexico	2.34	3.11	0.79	4.04
	China+	2.57	3.50	0.62	11.33
	Western Europe	1.44	3.00	0.71	3.76
	Russia+	1.61	4.68	0.64	16.92
	Global	0.94	2.14	0.50	6.90
Aggregate Industrial Electricity Demand	United States	0.44	1.30	0.79	10.32
	Mexico	1.09	2.61	0.62	5.60
	China+	0.88	3.39	0.30	17.06
	Western Europe	0.62	1.78	0.78	6.40
	Russia+	1.10	3.95	0.77	23.36
	Global	0.26	1.54	0.36	14.40
Aggregate Consumer Electricity Demand	United States	1.95	3.38	0.98	11.05
	Mexico	1.89	3.41	0.89	4.79
	China+	3.41	7.38	0.70	26.51
	Western Europe	1.76	3.26	0.99	7.29
	Russia+	2.30	4.30	0.78	12.36
	Global	0.80	2.18	0.68	13.09

forecasts. The t-tests for other bootstrap statistics, including the mean standard deviation and mean coefficient of variation, suggest that at least 30 model forecasts are necessary to replicate the full forecast set. Although replicating the characteristics of the mean model response can be done with few forecasts, the F-test for equality in the variances of the mean bootstrap resamples and the KS-test for equality of the empirical probability distribution functions indicate that larger forecast sets are necessary to fully capture the effect of uncertainty in the share parameters. As shown in the lower panel of Figure 5, the p-value of the F-test is greater than 0.05 when the subsample size is approximately 8,200 and larger. The KS-test has a p-value greater than 0.05 at a sample size of 7,500 and larger, indicating that we require at least 75% of the forecasts in order to be 95% confident that we can assume distributional equality with the full set of forecasts. Comparatively, the bootstrap results for Chinese gross domestic product suggest that approximately 6,700 forecasts are required to capture the variance in the bootstrap mean of the full set; however, slightly fewer forecasts are necessary to attain distributional equality with the full set. The KS-test indicates that approximately 65% of the full set are required in order to be 95% confident of distributional equality for Chinese gross domestic product. Thus our previous observation that variables with smaller coefficient of variation require fewer simulations to characterize model response is substantiated with the bootstrap results. Smaller subsamples are require to capture the variance and distributional aspects of Chinese gross domestic product ($\%c_v = 4.15$) versus Chinese consumer electric demand ($\%c_v = 7.38$).

We also estimated the bias in the model under uncertainty in the expenditure parameters by comparing the gross domestic product averaged over the full set and the mean model forecast, where all expenditures and elasticities were set to their mean values. This bias proved to be minimal: at most 0.35% different from the mean model forecast. Furthermore, it remained relatively constant over time. A similar analysis for China consumer electric demand also showed minimal bias of at most 0.25% in 2063 between the full set and mean sample.

4.2 Sensitivity to Elasticity Uncertainty

Model response to uncertainty in the elasticity of substitution parameters differed significantly from the response to expenditure uncertainty. The reasons are twofold. First, the elasticity of substitution does not play a significant role in the base year. In particular, the solution to the base-year model has unit inputs, outputs, and prices due to the choice of share parameters, and no substitutions are made. The forecast variables in subsequent years, however, are highly dependent on price fluctuations between commodities, and the elasticity of substitution parameters have a significant impact. Second, since we assume the elasticity of substitution parameters are the same across regions, attenuation through aggregating to global scales does not occur.

Since the model requires several iterations from the base year until it stabilizes with respect to the elasticities of substitution, we assess the response to uncertainty in these parameters by examining forecast variables in 2010 and 2063, rather than 2004 and 2063 as we did in examining the share parameters. In Table 4 the coefficient of variation and sample correlation between 2010 and 2063 of several forecast variables summarize the 4,978 viable model forecasts performed to assess sensitivity to uncertainty in the 70 elasticity of substitution parameters. Forecast variables tend to be less sensitive to uncertainty in the elasticities of substitution in (L) regions because relative price fluctuations in these economies are less pronounced than for (H) and (M) regions.

The global gross domestic product has a coefficient of variation of only approximately 2.5% in 2063, while other global aggregates have substantially higher coefficients of variation. Global CO_2 emissions, for example, have a coefficient of variation of approximately 20% in 2063. Many factors contribute to this observed difference in sensitivity between global gross domestic product

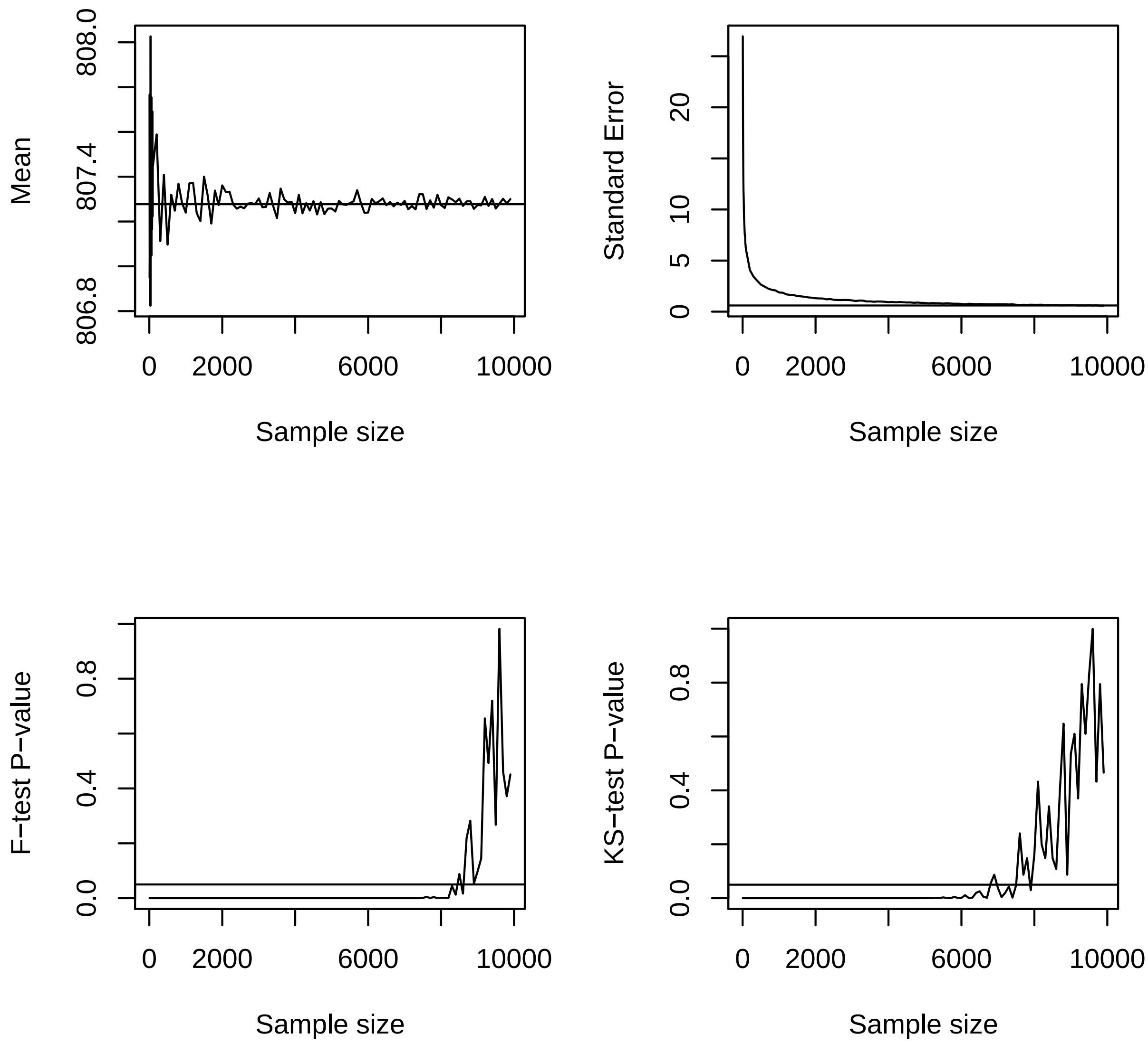

Figure 5: Bootstrap resampling of Chinese consumer demand for electricity in 2063 from the forecasts examining uncertainty in share parameters. Top left: bootstrap mean versus subsample size; the horizonal line is the bootstrap mean of the full set (807.3 Mtoe). Top right: bootstrap standard error of the mean versus subsample size; the horizonal line is the bootstrap standard error of the full set (0.6 Mtoe). Bottom left: F-test p-value for equality of the variance of the bootstrap mean; the horizonal line is the 95% confidence level (0.05). Bottom right: KS-test p-value for equality of distributions of the bootstrap mean; the horizonal line is the 95% confidence level (0.05).

Table 4: Coefficient of variation, sample correlation, and linear regression coefficients for select variables and regions over the 4,978 viable forecasts performed to assess sensitivity to uncertainty in the 70 elasticity of substitution parameters.

Variable	**Region**	$\%c_v$ 2010	$\%c_v$ 2063	r	β
Gross Domestic Product	United States	0.11	2.21	0.93	78.84
	Mexico	0.20	6.01	0.61	65.32
	China+	0.68	18.54	0.23	40.78
	Western Europe	0.32	4.35	0.49	21.66
	Russia+	0.64	6.60	0.33	18.60
	Global	0.14	2.57	0.60	45.50
Aggregate CO_2 Emissions	United States	0.39	14.39	0.75	64.82
	Mexico	0.33	16.08	0.88	107.93
	China+	0.43	23.85	0.19	41.83
	Western Europe	0.50	15.79	0.64	38.69
	Russia+	1.01	23.68	0.88	72.23
	Global	0.42	17.79	0.75	92.67
Steel and Iron Industry Revenue	United States	0.19	7.56	0.69	88.37
	Mexico	0.34	11.38	0.57	59.87
	China+	0.69	16.89	0.25	49.89
	Western Europe	0.77	13.19	0.69	26.98
	Russia+	3.46	10.06	0.47	8.01
	Global	0.19	12.44	0.47	140.07
Aggregate Industrial Electricity Demand	United States	0.51	4.28	0.77	23.19
	Mexico	0.24	11.43	0.82	119.96
	China+	0.66	18.59	0.12	30.80
	Western Europe	0.50	3.16	0.68	10.72
	Russia+	0.87	16.03	0.22	23.27
	Global	0.33	11.55	0.29	50.53
Aggregate Consumer Electricity Demand	United States	0.18	8.50	0.82	189.84
	Mexico	0.25	17.27	0.88	153.10
	China+	0.67	23.09	0.35	129.52
	Western Europe	0.25	6.38	0.38	33.27
	Russia+	1.30	29.22	0.94	120.76
	Global	0.40	16.33	0.94	206.01

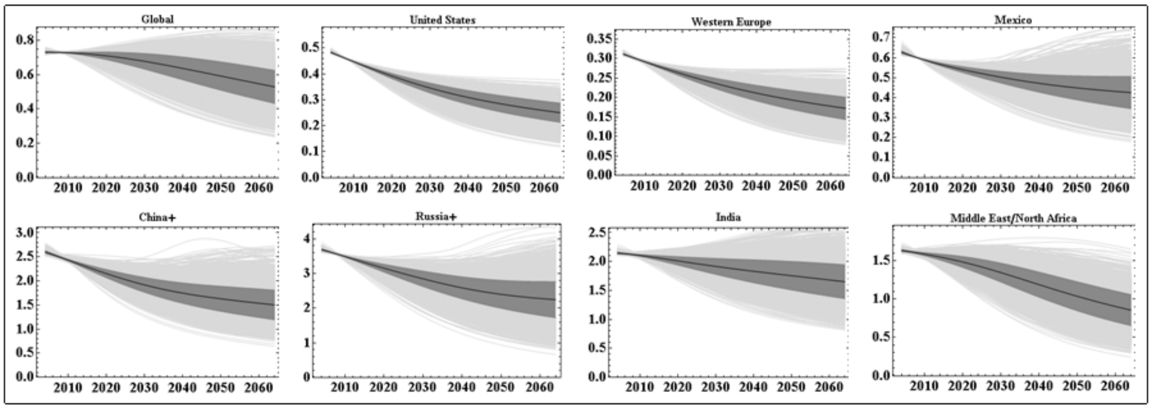

Figure 6: Carbon intensity forecast for the world and 7 of 16 model regions in kg CO_2 equivalent emissions per 2004 USD of gross domestic product.

and CO_2 emissions. One is that the gross domestic product is an aggregate of many variables and is dominated by more stable economies, whereas CO_2 emissions are dominated by less stable economies. Further, the year 2063 gross domestic product and CO_2 emissions are not correlated for the globe ($r = -0.007$) or for (L) regions such as the United States ($r = -0.054$) but have a positive correlation for (H) and (M) regions such as China ($r = 0.597$). These differences likely account for most of the discrepancy in sensitivity between these two variables.

Figure 6 shows the model forecast for global and regional carbon intensity (CO_2 emissions per unit gross domestic product) over for the 4,978 samples. Regional carbon intensities had uncertainties largely comparable to the uncertainty in emissions themselves. In other words, we find little correlation between the gross domestic product and emissions response variables for nearly all regions.

The bootstrap resamples of this set indicate that uncertainty in the elasticities of substitution results in a larger amount of variability in the model forecasts. Again we examine Chinese consumer electric demand in 2063 and find that the mean depicted in Figure 7 is easily recoverable from as few as 5 forecasts. The bootstrap standard error is very large (2.6 Mtoe) compared to that observed for the share parameter uncertainty set (0.6 Mtoe). The F-test for equality in variances indicates that at least 4,100 forecasts were required to replicate the full set of forecasts with 95% confidence. The KS-test indicated that approximately 4,200, or 84% of the full set would be necessary to obtain the same distributional characteristics with 95% confidence. The bootstrap resamples of Chinese gross domestic product show similar results. As we found with the uncertainty in the share parameters, slightly fewer forecasts are required to ensure variance and distributional equality with the full set at 95% confidence. This result is also seen in the coefficient variation, which is larger for consumer electric demand than for gross domestic product.

For the smaller set of forecasts where we examine the effect of uncertainty in the 16 Armington international trade elasticities of substitution, we allow the model to stabilize to 2020. The coefficient of variation and sample correlation between 2020 and 2063 of several forecast variables are summarized in Table 5. Forecast variables were most sensitive in regions and industries where international trade is essential. Relative to other regions, China displayed substantial sensitivity to uncertainty in the Armington elasticity parameters. Sizable sensitivities were also apparent for industry revenue variables that had a large revenue component from global trade, such as the steel and iron industry. While these sensitivities appear small relative to the sensitivities in our larger

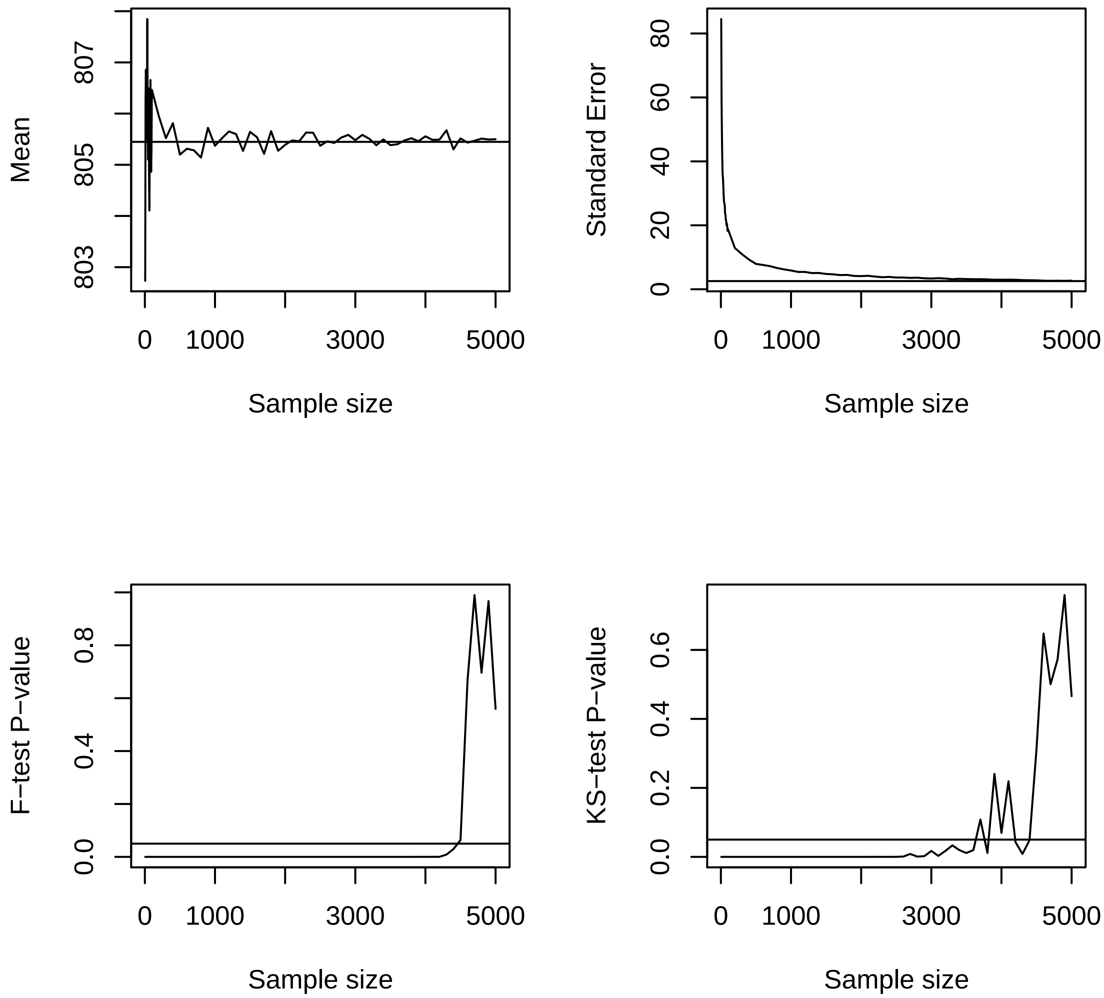

Figure 7: Bootstrap resampling of Chinese consumer demand for electricity in 2063 from the forecasts examining uncertainty in elasticities of substitution. Top left: bootstrap mean versus subsample size; the horizonal line is the bootstrap mean of full set (805.4 Mtoe). Top right: bootstrap standard error of the mean versus subsample size; the horizonal line is the bootstrap standard error of full set (2.6 Mtoe). Bottom left: F-test p-value for equality of the variance of the bootstrap mean; the horizonal line is the 95% confidence level (0.05). Bottom right: KS-test p-value for equality of distributions of the bootstrap mean; the horizonal line is the 95% confidence level (0.05).

set of elasticity of substitution parameters, they are vital when studying the impacts of mitigation policies on international trade. The bootstrap results summarized in Figure 8 indicate that there is less variability in Chinese consumer electric demand when we introduce uncertainty in the Armington set than when we introduce uncertainty in all elasticity parameters. Again there is no statistical difference in the bootstrap means between the subsample sets and the full set. In order to capture the variance, approximately 85% of the full set is required. The KS-test indicated that we would be 95% confident that the distribution of the full set can be characterized with approximately 650 forecasts.

Again, we estimated the bias in the model by comparing the gross domestic product averaged over the uncertainty sets and the mean model forecast, where all share parameters and elasticities are set to their mean values. For the 4,978 forecast set this bias remains near zero until 2020. For most regions it was less than 0.5%, but two notable exceptions are China and Africa, where the difference from the mean forecast is as much as 3% by 2063. For the 999 forecast set of Armington elasticities the bias remained near zero until 2040, and reached a maximum of only 0.4% by 2063.

5 Conclusions

We explored the sensitivity of the CIM-EARTH CGE model to uncertainty in the calibration dataset of expenditure parameters and the elasticities of substitution. Stark contrasts existed in the behavior of the model to these uncertainties, and our conclusions were highly dependent on the forecast variable being examined. For instance, the forecast variables having a smaller amount of initial uncertainty and higher level of aggregation displayed less sensitivity to expenditure uncertainty. The coefficient of variation was substantially smaller for global gross domestic product than for Chinese consumer electric demand. Furthermore, the average global gross domestic product over the 9,906 forecasts was within 0.2% of the mean sample, indicating that few samples would be necessary to characterize the mean response of the model. This finding was confirmed with the bootstrap results, which indicated that we could replicate the mean and standard deviation of the 9,906 forecast set with 99.9% confidence with as few as 100 model forecasts.

The effect of uncertainty in the elasticities of substitution was much more significant. The coefficient of variation for the 2063 Chinese consumer electric demand was 3 times larger than it was when we observed uncertainty in the expenditure data. We also found that the correlations between 2010 and 2063 were weaker for small-scale variables such as Chinese consumer electricity demand than their global aggregated counterparts. Furthermore, they were also weaker when we examined uncertainty in the elasticity parameters than they were for the expenditure uncertainty set.

Across all three uncertainty sets, the bootstrap results suggested that while we can confidently estimate the mean model response with few forecasts. To capture the variability and empirical probability distribution functions of the forecast variables, however, at least 75% of the total number of forecasts we conducted were necessary. Furthermore, our observation that variables with smaller coefficients of variation require fewer simulations to characterize model response was substantiated with the bootstrap results. That is, smaller subsamples were required to capture the variance and distributional aspects of Chinese gross domestic product ($\%c_v = 4.15$) versus Chinese consumer electric demand ($\%c_v = 7.38$).

Table 5: Coefficient of variation, sample correlation, and linear regression coefficients for select variables and regions over the 999 viable forecasts performed to assess sensitivity to uncertainty in the 16 Armington international trade elasticity parameters.

		$\%c_v$			
Variable	**Region**	2020	2063	r	β
Gross Domestic Product	United States	0.01	0.15	0.90	54.83
	Mexico	0.08	0.55	0.83	14.79
	China+	0.38	2.17	0.36	7.50
	Western Europe	0.05	0.33	0.22	3.47
	Russia+	0.23	1.00	0.70	10.18
	Global	0.05	0.55	0.33	10.85
Aggregate CO_2 Emissions	United States	0.05	0.54	0.31	5.85
	Mexico	0.23	0.38	0.31	1.00
	China+	0.20	3.00	0.38	14.15
	Western Europe	0.11	0.45	0.52	3.43
	Russia+	0.16	1.27	0.17	3.18
	Global	0.20	1.03	0.53	5.80
Steel and Iron Industry Revenue	United States	0.20	1.95	0.96	21.81
	Mexico	0.09	0.55	0.94	13.12
	China+	0.63	5.21	0.49	17.93
	Western Europe	0.26	1.31	0.97	8.65
	Russia+	0.36	1.24	0.38	4.48
	Global	0.16	2.98	0.45	25.82
Aggregate Industrial Electricity Demand	United States	0.02	0.27	0.87	28.35
	Mexico	0.03	0.29	0.41	9.25
	China+	0.26	3.80	0.55	37.90
	Western Europe	0.06	0.31	0.75	7.52
	Russia+	0.30	0.95	0.40	4.21
	Global	0.10	1.48	0.43	22.01
Aggregate Consumer Electricity Demand	United States	0.02	0.21	0.52	16.93
	Mexico	0.05	0.87	0.58	18.69
	China+	0.21	3.11	0.80	64.13
	Western Europe	0.02	0.19	0.51	10.55
	Russia+	0.08	1.14	0.53	26.76
	Global	0.04	1.03	0.86	81.87

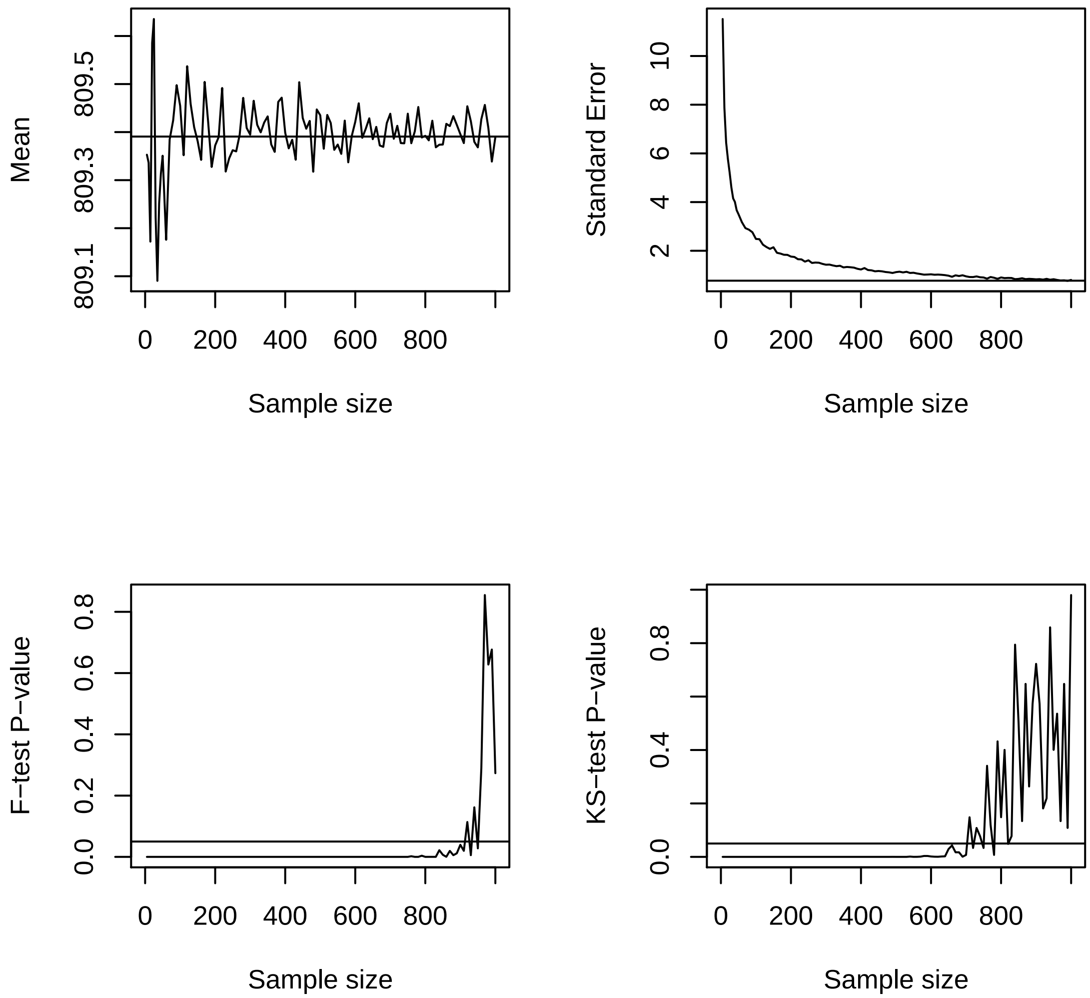

Figure 8: Bootstrap resampling of Chinese consumer demand for electricity in 2063 from the forecasts examining uncertainty in Armington elasticities of substitution. Top left: bootstrap mean versus subsample size; the horizonal line is the bootstrap mean of full set (809.4 Mtoe). Top right: bootstrap standard error of the mean versus subsample size; the horizonal line is the bootstrap standard error of full set (0.8 Mtoe). Bottom left: F-test p-value for equality of the variance of the bootstrap mean; the horizonal line is the 95% confidence level (0.05). Bottom right: KS-test p-value for equality of distributions of the bootstrap mean; the horizonal line is the 95% confidence level (0.05).

Acknowledgments

We thank Kenneth Judd for his advice during the preparation of this manuscript and Michael Wilde and Allan Espinosa for their help using the Swift parallel scripting system. Batches of forecasts were performed on the TeraGrid (TG) and Open Science Grid (OSG) machines, including Firefly (University of Nebraska, OSG), QueenBee (Louisiana State University, TG), Ranger (Argonne National Laboratory, TG), TeraPort (University of Chicago Computation Institute, OSG), and Texas Advanced Computing Center (University of Texas, TG).

A Parameter Distributions

Figure 9 compares the parameter distributions from several studies for the elasticity of substitution between capital and labor by industry. Given the sizable discrepancies in the estimates of means and standard deviations, we use Gaussian parameter distributions with relatively standard mean values and standard deviations set to 20% of the mean.

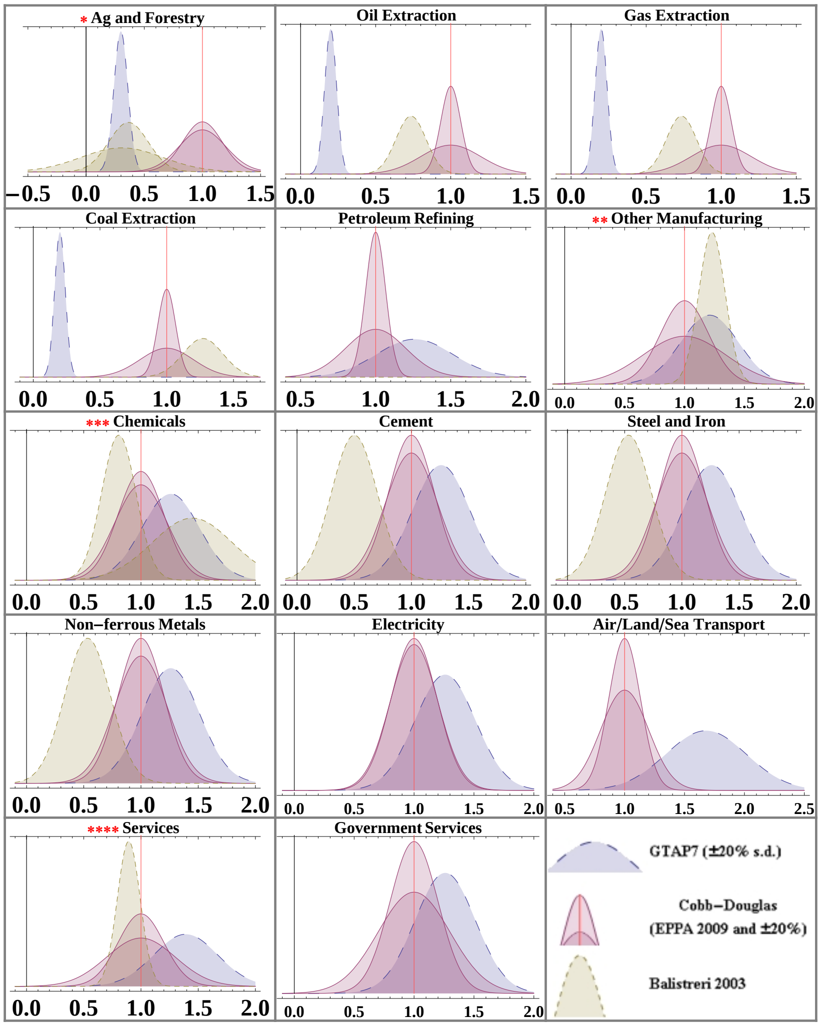

Figure 9: A comparison of parameter distributions for σ_{KL} for a variety of industries. The red line at σ=1 denotes the Cobb-Douglas point, and the black line at σ=0 denotes the Leontief or fixed-coefficients point. Some aggregate industries have multiple estimates from [3] that are relevant: * Balistreri et al. estimate two agriculture-related industries: "farms" and "agriculture and forestry services." ** Balistreri et al. estimate many industries relevant to generic manufacturing; the estimate for aggregated manufacturing and mining is shown here. *** Balistreri et al. estimate two chemical-related industries: "rubber and misc. plastic products" and "chemicals and allied products." **** Balistreri et al. estimate only one service industry: "construction services," which should not be taken as representative of aggregated services.

References

[1] D. Abler, Rodriguez, A.G., and J. Shortle, *Parameter uncertainty in CGE modeling of the environmental impacts of economic policies*, Environmental and Resource Economics, 14 (1999), pp. 75–94.

[2] M. H. Babiker, J. Reilly, M. Mayer, R. S. Eckaus, I. S. Wing, and R. C. Hyman, *The MIT Emissions Prediction and Policy Analysis (EPPA) Model: Revisions, Sensitivities, and Comparisons of Results*, Technical Report 71, MIT Joint Program Report Series, 2001.

[3] E. J. Balistreri, C. A. McDaniel, and E. V. Wong, *An Estimation of U.S. Industry-Level Capital-Labor Substitution*, Computational Economics 0303001, EconWPA, March 2003.

[4] C. Ballard, D. Fullerton, J. B. Shoven, and J. Whalley, *A General Equilibrium Model for Tax Policy Evaluation*, National Bureau of Economic Research Monograph, The University of Chicago Press, 1985.

[5] M. Berrittella, A. Bigano, R. Roson, and R. S. J. Tol, *A general equilibrium analysis of climate change impacts on tourism*, Tourism Management, 27 (2006), pp. 913–924.

[6] M. Berrittella, A. Y. Hoekstra, K. Rehdanz, R. Roson, and R. S. J. Tol, *The economic impact of restricted water supply: A computable general equilibrium analysis*, Water Research, 41 (2007), pp. 1799–1813.

[7] C. Boehringer, T. F. Rutherford, and W. Wiegard, *Computable General Equilibrium Analysis: Opening a Black Box*, Discussion Paper 03-56, ZEW, 2003.

[8] F. Bosello, R. Roson, and R. S. J. Tol, *Economy-wide estimates of the implications of climate change: Human health*, Ecological Economics, 58 (2006), pp. 579–591.

[9] ——, *Economy-wide estimates of the implications of climate change: Sea level rise*, Environmental and Resource Economics, 37 (2007), pp. 549–571.

[10] ——, *Economy-wide estimates of the implications of climate change – a rejoinder*, Ecological Economics, 66 (2008), pp. 14–15.

[11] J.-M. Burniaux and T. Truong, *GTAP-E: An Energy-Environmental Version of the GTAP Model*, GTAP Technical Papers 923, Center for Global Trade Analysis, Department of Agricultural Economics, Purdue University, 2002.

[12] S. P. Dirkse and M. C. Ferris, *The PATH solver: A non-monotone stabilization scheme for mixed complementarity problems*, Optimization Methods and Software, 5 (1995), pp. 123–156.

[13] B. Effron and R. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, 1993.

[14] J. Elliott, I. Foster, K. Judd, E. Moyer, and T. Munson, *CIM-EARTH: Community Integrated Model of Economic and Resource Trajectories for Humankind*, Technical Memorandum ANL/MCS-TM-307 Version 0.1, MCS Division, Argonne National Laboratory, 2009.

[15] M. C. Ferris and T. Munson, *Interfaces to PATH 3.0: Design, implementation and usage*, Computational Optimization and Applications, 12 (1999), pp. 207–227.

[16] ——, *GAMS/PATH User Guide: Version 4.3*, Department of Computer Sciences, University of Wisconsin, Madison, 2000.

[17] R. Fourer, D. M. Gay, and B. W. Kernighan, *AMPL: A Modeling Language for Mathematical Programming*, Brooks/Cole–Thomson Learning, Pacific Grove, California, second ed., 2003.

[18] R. Gerlagh and B. Van der Zwaan, *A sensitivity analysis of timing and costs of greenhouse gas emission reductions*, Climatic Change, 65 (2004), pp. 39–71.

[19] V. Ginsburgh and M. Keyzer, *The Structure of Applied General Equilibrium Models*, The MIT Press, Cambridge, 1997.

[20] B. N. Gopalakrishnan and T. L. Walmsley, eds., *Global Trade, Assistance, and Production: The GTAP 7 Data Base*, Global Trade Analysis Center, Department of Agricultural Economics, Purdue University, 2008.

[21] T. Hertel, D. Hummels, M. Ivanic, and R. Keeney, *How confident can we be of CGE-based assessments of free trade agreements?*, Economic Modelling, 24 (2007), pp. 611–635.

[22] D. W. Jorgenson, *Growth, Volume 1: Econometric General Equilibrium Modeling*, MIT Press, 1998.

[23] D. Kendrick, *Stochastic control for economic models: Past, present and the paths ahead*, Journal of Economic Dynamics and Control, 29 (2005), pp. 3–30.

[24] S. Kim, *Uncertainty, political preferences, and stabilization: Stochastic control using dynamic CGE models*, Computational Economics, 24 (2004), pp. 97–116.

[25] J. Liu, C. Arndt, and T. Hertel, *Parameter estimation and measures of fit in a global, general equilibrium model*, Journal of Economic Integration, 19 (2004), pp. 626–649.

[26] H. E. Scarf and J. B. Shoven, *Applied General Equilibrium Analysis*, Cambridge University Press, Cambridge, UK, 1984.

[27] A. Sokolov, C. Schlosser, S. Dutkiewicz, S. Paltsev, D. Kicklighter, H. Jacoby, R. Prinn, C. Forest, J. Reilly, C. Wang, B. Felzer, M. Sarofim, J. Scott, P. Stone, J. Melillo, and J. Cohen, *The MIT Integrated Global System Model (IGSM) Version 2: Model Description and Baseline Evaluation*, Technical Report 124, MIT Joint Program Report Series, 2005.

[28] A. P. Sokolov, P. H. Stone, C. E. Forest, R. Prinn, M. C. Sarofim, M. Webster, S. Paltsev, C. A. Schlosser, D. Kicklighter, S. Dutkiewicz, J. Reilly, C. Wang, B. Felzer, J. M. Melillo, and H. D. Jacoby, *Probabilistic forecast for twenty-first-century climate based on uncertainties in emissions (without policy) and climate parameters*, Journal of Climate, 22 (2009), pp. 5175–5204.

[29] I. Sue Wing, *Computable General Equilibrium Models and Their Use in Economy-Wide Policy Analysis*, Technical Note 6, Joint Program on the Science and Policy of Global Change, 2004.

[30] R. Tol and S. Fankhauser, *On the representation of impact in integrated assessment models of climate change*, Environmental Modeling and Assessment, 3 (1998), pp. 63–74.

[31] M. Webster, S. Paltsev, J. Parsons, J. Reilly, and H. Jacoby, *Uncertainty in Greenhouse Emissions and Costs of Atmospheric Stabilization*, Technical Report 165, MIT Joint Program Report Series, 2008.

[32] M. Wilde, I. Foster, K. Iskra, P. Beckman, Z. Zhang, A. Espinosa, M. Hategan, B. Clifford, and I. Raicu, *Parallel scripting for applications at the petascale and beyond*, Computer, 42 (2009), pp. 50–60.

[33] Y. Zhao, M. Hategan, B. Clifford, I. Foster, G. von Laszewski, V. Nefedova, I. Raicu, T. Stef-Praun, and M. Wilde, *Swift: Fast, Reliable, Loosely Coupled Parallel Computation*, in 1st IEEE International Workshop on Scientific Workflows, 2007, pp. 199–206.